ON DERIVATIVES OF GRAPHON PARAMETERS

LÁSZLÓ MIKLÓS LOVÁSZ AND YUFEI ZHAO

Abstract. We give a short elementary proof of the main theorem in the paper "Differential calculus on graphon space" by Diao et al. (2015) [2], which says that any graphon parameters whose $(N+1)$-th derivatives all vanish must be a linear combination of homomorphism densities $t(H,-)$ over graphs H on at most N edges.

Let $\mathcal{W} \subset L^\infty([0,1]^2,\mathbb{R})$ denote the set of bounded symmetric measurable functions $f\colon [0,1]^2 \to \mathbb{R}$ (here symmetric means $f(x,y) = f(y,x)$ for all x,y). Let $\mathcal{W}_{[0,1]} \subset \mathcal{W}$ denote those functions in $\mathcal{W}$ taking values in $[0,1]$. Such functions, known as *graphons*, are central to the theory of graph limits [3], an exciting and active research area giving an analytic perspective towards graph theory.

In [2], the authors systematically study the local structure of differentiable graphon parameters. They develop the theory of consistency constraints for multilinear functionals on graphon space, and as a consequence, obtain the result (Theorem 1 below) that is the graphon analog of the following basic fact from calculus: the set of functions whose $(N+1)$-th derivatives all vanish identically is precisely the set of polynomials of degree at most N. For graphons, homomorphism densities $t(H,-)$ play the role of monomials: they generate a ring of smooth functions that separate points and they have the property of vanishing higher derivatives as in Theorem 1. In this short note, we follow a more direct route to prove their result. Our proof avoids the technicalities of the approach in [2].

We begin with some definitions. The space $\mathcal{W}$ is equipped with the *cut norm*

$$\|f\|_\square := \sup_{\text{measurable } S,T\subseteq[0,1]} \left| \int_{S\times T} f(x,y)\,dxdy \right|.$$

Given $g \in \mathcal{W}$, and a measure-preserving map $\phi\colon [0,1] \to [0,1]$, we define $g^\phi(x,y) := g(\phi(x),\phi(y))$. The *cut distance* on $\mathcal{W}$ is defined by $\delta_\square(f,g) := \inf_\phi \left\| f - g^\phi \right\|_\square$ where ϕ ranges over all such measure-preserving maps. Let $\sim$ denote the equivalence relations in $\mathcal{W}$ defined by $f \sim g \Leftrightarrow \delta_\square(f,g) = 0$. It is known that $(\mathcal{W}_{[0,1]}/\sim, \delta_\square)$ is a compact metric space [4].

Functions $F\colon \mathcal{W}_{[0,1]}/\sim \to \mathbb{R}$ are called *class functions* (we import this terminology from [2]; the term *graphon parameter* is also used in the literature). Class functions that are continuous with respect to the cut distance play an important role in graph parameter/property testing [1, 5].

Define the *admissible directions at* $f \in \mathcal{W}_{[0,1]}$ as

$$\mathrm{Adm}(f) := \{g \in \mathcal{W} : f + \epsilon g \in \mathcal{W}_{[0,1]} \text{ for some } \epsilon > 0\}.$$

The *Gâteaux derivative of* F *at* $f \in \mathcal{W}_{[0,1]}$ *in the direction* $g \in \mathrm{Adm}(f)$ is defined by (if it exists)

$$dF(f;g) := \lim_{\lambda\to 0^+} \frac{1}{\lambda}(F(f+\lambda g) - F(f)).$$

Higher mixed Gâteaux derivatives are defined iteratively: $d^{N+1}F(f;g_1,\dots,g_{N+1})$ is defined to be the Gâteaux derivative of $d^N F(-;g_1,\dots,g_N)$ at f in the direction g_{N+1}, if this limit exists.

Let $\mathcal{H}_n$ denote the isomorphism classes of multi-graphs with n edges, no isolated vertices, and no self-loops but possible multi-edges. Also let $\mathcal{H}_{\le n} := \bigcup_{j\le n} \mathcal{H}_j$ and $\mathcal{H} := \bigcup_{j\in\mathbb{N}} \mathcal{H}_j$.

Y. Zhao was supported by a Microsoft Research PhD Fellowship.

For any $H \in \mathcal{H}$, and any $f \in \mathcal{W}$, we define the homomorphism density

$$t(H,f) := \int_{[0,1]^{V(H)}} \prod_{ij \in E(H)} f(x_i,x_j) \prod_{i \in V(H)} dx_i,$$

where $E(H)$ is the multi-set of edges of H. For example, when H consists of two vertices and two parallel edges between them, $t(H,W) = \int_{[0,1]^2} W(x,y)^2\,dxdy$.

Here is the main result of [2].

Theorem 1 (Diao, Guillot, Khare, Rajaratnam [2, Theorem 1.4])**.** *Let $F\colon \mathcal{W}_{[0,1]} \to \mathbb{R}$ be a class function which is continuous with respect to the L^1 norm and $N+1$ times Gâteaux differentiable for some $N \geq 0$. Then F satisfies*

$$d^{N+1}F(f; g_1,\dots,g_{N+1}) = 0, \qquad \forall f \in \mathcal{W}_{[0,1]},\ g_1,\dots,g_{N+1} \in \mathrm{Adm}(f),$$

if and only if there exist constants c_H such that

$$F(f) = \sum_{H \in \mathcal{H}_{\leq N}} c_H t(H,f). \tag{1}$$

Moreover, the constants c_H are unique. If in addition F is continuous with respect to the cut norm, then $c_H = 0$ if $H \in \mathcal{H}_{\leq N}$ is not a simple graph.

The "if" direction is simple. From the definition, we can see that $t(H, f + \lambda_1 g_1 + \cdots + \lambda_{N+1} g_{N+1})$ expands into a polynomial in $\lambda_1,\dots,\lambda_{N+1}$ of total degree at most $|E(H)| \leq N$, which clearly implies that its derivative with respect to $d\lambda_1 d\lambda_2 \dots d\lambda_{N+1}$ vanishes identically. Thus any F of the form (1) satisfies $d^{N+1}F \equiv 0$ (and is L^1-continuous).

For the "only if" direction, we first give a sketch. When the domain of F is restricted to graphons that correspond to edge-weighted graphs on n vertices, F is simply a function on $\binom{n}{2}$ real variables. So the vanishing of its $(N+1)$-th order derivatives implies that it is a polynomial of degree at most N. From these polynomials we can recover the coefficients of $t(H,-)$. Weighted graphs on finitely many vertices correspond to graphons that are step functions, and they are dense in $\mathcal{W}_{[0,1]}$ with respect to the L^1 norm, so the claim follows by continuity.

Now come the details. Let $\mathcal{M}_n$ denote the set of symmetric $n \times n$ matrices $a = (a_{i,j})$ with zeros on the diagonal ($a_{i,i} = 0$), and let $\mathcal{M}_{n,[0,1]} \subset \mathcal{M}_n$ be the matrices with entries in $[0,1]$. We view elements of $\mathcal{M}_n$ as edge-weighted complete graphs on n labeled vertices. For $a, b \in \mathcal{M}_n$, we write $a \sim b$ if a can be obtained from b by a permutation of the vertex labels. We define class functions and Gâteaux derivatives for $\mathcal{M}_n$ analogously to how they are defined for $\mathcal{W}$. Write $[n] := \{1,\dots,n\}$. For any $a \in \mathcal{M}_n$ and $H \in \mathcal{H}$ (assume that $V(H) = \{1,\dots,|V(H)|\}$), define

$$t(H,a) = \frac{1}{n^{|V(H)|}} \sum_{v_1,\dots,v_{|V(H)|} \in [n]} \prod_{ij \in E(H)} a_{v_i,v_j}. \tag{2}$$

There is a natural embedding $\mathcal{M}_n \hookrightarrow \mathcal{W}$, identifying $a \in \mathcal{M}_n$ with $f_a \in \mathcal{W}$ given by $f_a(x,y) = a_{\lceil nx \rceil, \lceil ny \rceil}$ (and $f_a(x,y) = 0$ if x or y is 0). All previous notions are consistent with the identification.

Note that $t(H,a)$ is a degree $|E(H)|$ polynomial in $a_{i,j}$, $1 \leq i < j \leq n$ (recall that a was symmetric, so $a_{i,j} = a_{j,i}$). Write $(n)_k := n(n-1)\cdots(n-k+1)$ and define

$$t^{\mathrm{inj}}(H,a) = \frac{1}{(n)_{|V(H)|}} \sum_{\text{distinct } v_1,\dots,v_{|V(H)|} \in [n]} \prod_{ij \in E(H)} a_{v_i,v_j}. \tag{3}$$

For each fixed H and $n \geq |V(H)|$, $t(H,a)$ equals a nonzero multiple of $t^{\mathrm{inj}}(H,a)$ plus a linear combination of various $t^{\mathrm{inj}}(H',a)$ with $|E(H')| = |E(H)|$ and $|V(H')| < |V(H)|$ (essentially recording the different ways that $v_1,\dots,v_{|V(H)|}$ can fail to be distinct in the summation for $t(H,a)$). It follows that $(t^{\mathrm{inj}}(H,-) : H \in \mathcal{H}_N)$ can be transformed into $(t(H,-) : H \in \mathcal{H}_N)$ via a lower triangular

matrix with positive diagonal entries (when $\mathcal{H}_N$ is sorted by the number of vertices), and vice versa (since such matrices are invertible).

Let $\mathcal{H}_d^{(n)}$ consist of those $H \in \mathcal{H}_d$ with at most n vertices. The main observation we need to make is the following lemma:

Lemma 2. *If a class function $F\colon \mathcal{M}_{n,[0,1]} \to \mathbb{R}$ is a homogeneous polynomial of degree d, then we can write $F = \sum_{H\in\mathcal{H}_d^{(n)}} c_H t^{\mathrm{inj}}(H,-)$ for some $c_H \in \mathbb{R}$, in a unique way.*

Proof. Since F is a class function, the coefficient of the monomial $a_{i_1,j_1}\dots a_{i_d,j_d}$ is equal to the coefficient of $a_{\sigma(i_1),\sigma(j_1)}\dots a_{\sigma(i_d),\sigma(j_d)}$ for all permutations σ of $[n]$. Observe that the polynomial $\sum_{\sigma\in S_n} a_{\sigma(i_1),\sigma(j_1)}\dots a_{\sigma(i_d),\sigma(j_d)}$ is a multiple of $t^{\mathrm{inj}}(H,a)$ for the multigraph H whose multi-set of edges is given by $E(H) = \{i_1j_1,\dots,i_dj_d\}$. For distinct H and H', the set of monomials that appear in $t^{\mathrm{inj}}(H,a)$ and $t^{\mathrm{inj}}(H',a)$ are disjoint. Thus, we have a direct correspondence between linear combinations of $t^{\mathrm{inj}}(H,-)$ for $H \in \mathcal{H}_d^{(n)}$ and polynomials of degree d. $\square$

In particular, this lemma implies the following:

Lemma 3. *The elements of $\{t(H,-) : H \in \mathcal{H}_{\le N}\}$ are linearly independent as functions on $\mathcal{M}_{n,[0,1]}$ whenever $n \ge 2N$.*

Proof. If $n \ge 2N$, then any graph H with at most N edges and no isolated vertices has at most $2N$ vertices. Thus the polynomials $\{t^{\mathrm{inj}}(H,-),\ H \in \mathcal{H}_{\le N}\}$ are linearly independent. By the linear relations between $\{t(H,-)\}$ and $\{t^{\mathrm{inj}}(H,-)\}$, it follows that $\{t(H,-) : H \in \mathcal{H}_{\le N}\}$ is linearly independent as well. $\square$

Lemma 4. *If $F\colon \mathcal{M}_{n,[0,1]} \to \mathbb{R}$ is a class function whose $(N+1)$-th derivatives vanish everywhere, then $F = \sum_{H\in\mathcal{H}_{\le N}} c_H t(H,-)$ for some $c_H \in \mathbb{R}$. If $n \ge 2N$, the values c_H are uniquely determined.*

Proof. Note that $\mathcal{M}_{n,[0,1]}$ is a subset of a finite dimensional vector space, which means F is a function of $\binom{n}{2}$ real variables, and its Gâteaux derivatives are just the usual partial derivatives. So if the $(N+1)$-th derivatives of F all vanish, then F must be a polynomial of degree at most N. By Lemma 2, F lies in the span of $t^{\mathrm{inj}}(H,-)$, $H \in \mathcal{H}_{\le N}$, and hence it lies in the span of $t(H,-)$, $H \in \mathcal{H}_{\le N}$. By Lemma 3, if $n \ge 2N$, the functions $t(H,-)$ are linearly independent, so the values c_H are unique. $\square$

Now we prove the "only if" direction of Theorem 1. By embedding $\mathcal{M}_n \hookrightarrow \mathcal{W}$, the hypothesis $d^{N+1}F \equiv 0$ on $\mathcal{M}_{n,[0,1]}$ implies, by Lemma 4, that $F = \sum_{H\in\mathcal{H}_{\le N}} c_H^{(n)} t(H,-)$ on $\mathcal{M}_{n,[0,1]}$ for some $c_H^{(n)}$, uniquely if $n \ge 2N$. For any $m,n \ge 2N$ with $m/n \in \mathbb{N}$, the image of $\mathcal{M}_n$ in $\mathcal{W}$ is contained in the image of $\mathcal{M}_m$. Since $F = \sum_{H\in\mathcal{H}_{\le N}} c_H^{(m)} t(H,-)$ on $\mathcal{M}_m$, restricting to $\mathcal{M}_n$, we see that $c_H^{(n)} = c_H^{(m)}$ for all $H \in \mathcal{H}_{\le N}$. It then follows that for any $n,n' \ge 2N$, $c_H^{(n)} = c_H^{(nn')} = c_H^{(n')}$, so there is some c_H so that $c_H^{(n)} = c_H$ for all $n \ge 2N$.

It follows that $F = \sum_{H\in\mathcal{H}_{\le N}} c_H t(H,-)$ on $\bigcup_{n\in\mathbb{N}} \mathcal{M}_{n,[0,1]}$, whose image is dense in $\mathcal{W}_{[0,1]}$ with respect to the L^1 norm. As both sides of the equation are continuous with respect to the L^1 norm, the equality holds in all of $\mathcal{W}_{[0,1]}$. The uniqueness of the constants c_H follows from Lemma 4.

The proof of the final claim in Theorem 1 is reproduced here from [2] for completeness. Suppose F is continuous with respect to the cut norm. Then

$$F(f) = \sum_{H\in\mathcal{H}_{\le N}} c_H t(H^{\mathrm{simple}}, f) \tag{4}$$

where H^{simple} is the simple graph obtained from H by replacing any multi-edge by a single edge between the same pair of vertices. Indeed, (4) holds for $\{0,1\}$-valued f since $t(H^{\mathrm{simple}}, f) = t(H, f)$

for all $\{0,1\}$-valued f. Since the set of $\{0,1\}$-valued graphons is dense in $\mathcal{W}_{[0,1]}$ with respect to cut distance, and both sides of (4) are continuous in f with respect to cut distance, (4) holds on all of $\mathcal{W}_{[0,1]}$. Thus only simple graphs are needed in the summation for F.

References

[1] C. Borgs, J. T. Chayes, L. Lovász, V. T. Sós, and K. Vesztergombi, *Convergent sequences of dense graphs. I. Subgraph frequencies, metric properties and testing*, Adv. Math. **219** (2008), 1801–1851.

[2] P. Diao, D. Guillot, A. Khare, and B. Rajaratnam, *Differential calculus on graphon space,* J. Combin. Theory Ser. A **133** (2015), 183–227.

[3] L. Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.

[4] L. Lovász and B. Szegedy, *Szemerédi's lemma for the analyst*, Geom. Funct. Anal. **17** (2007), 252–270.

[5] L. Lovász and B. Szegedy, *Testing properties of graphs and functions*, Israel J. Math. **178** (2010), 113–156.

Department of Mathematics, MIT, Cambridge, MA 02139, United States
E-mail address: `lmlovasz@math.mit.edu`

Mathematical Institute, University of Oxford, Oxford OX2 6GG, United Kingdom
E-mail address: `yufei.zhao@maths.ox.ac.uk`